

# A new level-dependent coarsegrid correction scheme for indefinite Helmholtz problems

Siegfried Cools, Bram Reys, Wim Vanroose\*

*Department of Mathematics and Computer Science, University of Antwerp, Middelheimlaan 1, 2020 Antwerp, Belgium  
Intel ExaScience Lab, Kapeldreef 75, B-3001 Leuven, Belgium (B. Reys).*

## SUMMARY

In this paper we construct and analyse a level-dependent coarsegrid correction scheme for indefinite Helmholtz problems. This adapted multigrid method is capable of solving the Helmholtz equation on the finest grid using a series of multigrid cycles with a grid-dependent complex shift, leading to a stable correction scheme on all levels. It is rigorously shown that the adaptation of the complex shift throughout the multigrid cycle maintains the functionality of the two-grid correction scheme, as no smooth modes are amplified in or added to the error. In addition, a sufficiently smoothing relaxation scheme should be applied to ensure damping of the oscillatory error components. Numerical experiments on various benchmark problems show the method to be competitive with or even outperform the current state-of-the-art multigrid-preconditioned Krylov methods, like e.g. CSL-preconditioned GMRES or BiCGStab. Copyright © 2010 John Wiley & Sons, Ltd.

Received ...

**KEY WORDS:** Helmholtz equation, indefinite systems, multigrid, level-dependent correction scheme, complex shifted Laplacian

## 1. INTRODUCTION

Originally introduced by Brandt in 1976 [8], the multigrid method is known to be a particularly fast and scalable solver for the large systems of equations arising from the discretisation of multi-dimensional Poisson or positive definite Helmholtz equations, see [34, 35, 11]. Using a negative shift however, the discretisation matrix becomes distinctly indefinite causing multigrid convergence to deteriorate. This break-down is due to near-to-zero eigenvalues in the operator spectrum on some intermediate level in the multigrid hierarchy, which destroy the functionality of the standard correction scheme, as shown by Elman, Ernst and O’leary in [14].

Motivated primarily by geophysical applications [29], rising interest in the development of fast solvers for the indefinite Helmholtz equation over the past decade has yielded a broad range of solution methods for this non-trivial problem, an overview of which can be found in [20]. Due to their wide range of applications on a variety of problems, preconditioned Krylov subspace methods are currently among the most popular Helmholtz solvers. The Krylov subspace method functions as an outer iteration and a direct (ILU) or iterative (multigrid) method is used to (approximately) solve the preconditioning system in every step. Being crucial to the performance of the governing Krylov method, significant research has been performed over the construction of a good preconditioner. Past and recent work includes the wave-ray approach [9], the idea of separation of variables [25], algebraic multilevel methods [7] and a transformation of the Helmholtz equation into an advection-diffusion-reaction problem [21]. The key ideas of the current paper shows some resemblances to the work done in [23].

\*Correspondence to: siegfried.cools@ua.ac.be, wim.vanroose@ua.ac.be

Another multigrid-based preconditioner was suggested in [14], which has led to the development of the so-called shifted Laplacian preconditioners. The first papers on these preconditioners are [4] and [22], in which a Laplace operator with a real shift was proposed as a preconditioner instead of the approximate inverse of the original Helmholtz operator. However, still being unable to efficiently solve very high wavenumber problems using this approach, the idea was reinvented and extended to the Complex Shifted Laplacian (CSL) by Erlangga, Vuik and Oosterlee in [17, 18] and analysed further in [19, 36]. Leading to satisfactory convergence and scalability results on highly indefinite problems, the complex shifting of the original problem operator was generalised in [15, 16, 2, 24].

In [27] it was shown that scaling the wavenumber by a complex value is equivalent to scaling the grid distance, thereby rotating the spectrum around its most negative real point instead of translating it down (or up) in the complex plane. The effects of this Complex Stretched Grid (CSG) transformation on MG-preconditioned Krylov solvers were analysed to some extent in [26]. The resulting preconditioner is a Helmholtz operator discretised on a complex-valued grid, which is proven to be particularly efficient when complex-valued grid distances are already used to implement advanced absorbing boundary conditions like ECS [1] or PML [6].

In this paper we aim to solve the indefinite Helmholtz equation using a new level-dependent multigrid scheme. The method adopts the key idea from CSL/CSG, which is plugged into the multigrid correction scheme, gradually shifting the eigenvalues away from the origin in the coarsening process. This results in a multigrid scheme which, contrarily to the above methods, can be applied *directly* as a solver to the Helmholtz equation, instead of merely suiting preconditioning purposes. Numerical results show the method to be competitive with (or even outperform) the current state-of-the-art MG-preconditioned Krylov solvers.

To complement this adapted correction scheme, a stable relaxation scheme with profound smoothing properties should be used to effectively damp the oscillatory error components. It is well-known that standard smoothers such as weighted-Jacobi or Gauss-Seidel do not necessarily resolve high-frequency modes and might even display divergent behaviour. Considering this observation we recommend using the more advanced GMRES(3) method as a smoother. First suggested as a replacement for a standard multigrid smoother in [14], it was recently shown in [28] and [12] that replacing the standard smoothers by GMRES(3) yields very satisfactory multigrid convergence results.

The paper is organised as follows. In Section 2 we briefly recall the properties of both the CSL and CSG preconditioners. Additionally, we show the stability and functionality of these schemes using the spectral analysis from [14]. In Section 3 we then define the new level-dependent correction scheme, in which we differentiate between a CSL- and CSG-based variant. The level-dependent scheme is analysed through a standard two-grid spectral analysis, emphasizing the complementary action of the correction scheme and the smoother. Numerical results supporting the theory are shown in Section 4, where the new level-dependent scheme is extensively tested on a variety of benchmark problems, ranging from the academic constant wavenumber model problem to more hard-to-solve quantum mechanical models featuring evanescent waves. Finally, along with a discussion on the subject, conclusions are drawn in Section 5.

## 2. OVERVIEW OF COMPLEX SHIFTED PRECONDITIONERS

It is the aim of this work to effectively solve the discretised Helmholtz system

$$(-\Delta^h - k^2)u^h = \chi^h, \quad \text{on } \Omega^h \subset \mathbb{R}^d, \quad (1)$$

with dimension  $d \geq 1$ , where  $\Delta$  is the Laplace operator,  $\chi$  represents the righthand side or *source term*, and  $k \in \mathbb{R}$  is known as the (possibly spatial dependent) *wavenumber* which might cause the system to become highly indefinite. Note that for the analysis we generally restrain ourselves to the one-dimensional problem formulation with Dirichlet boundary conditions and a constant wavenumber  $k$ , as is common practice in Helmholtz literature. The theoretical results obtained in this work can however readily be extended to more general problem settings in higher dimensions.

Successful experiments in higher spatial dimensions (see Section 4) will provide a solid foundation for this statement.

### 2.1. Two-grid spectral analysis

As a multigrid cycle is essentially a recursive embedding of two-grid correction schemes on consecutive coarser grids, flanked on both sides by a pre- and postsmoothing operator, an analysis of the two-grid scheme is crucial in understanding the full action of the multigrid method. The two-grid correction scheme for an error  $e^h \in \Omega^h$  is given by the matrix-operator (see [11], [34])

$$TG = (I - I_{2h}^h (A^{2h})^{-1} I_h^{2h} A^h) \quad (2)$$

where  $I$  is the identity matrix,  $A^h$  and  $A^{2h}$  are the fine- and coarsegrid discretisation matrices respectively,  $I_{2h}^h$  is the interpolation or prolongation operator, and  $I_h^{2h}$  is the restriction operator<sup>†</sup>. Note that any vector  $v^h \in \Omega^h$ , specifically the error  $e^h$ , can be written as a linear combination of the eigenvectors of  $A^h$

$$e^h = \sum_{l=1}^n a_l w_l^h, \quad (3)$$

$a_l \in \mathbb{R}$ , where it is known that the eigenvectors of  $A^h$  are given componentwise by

$$w_{l,j}^h = \sin\left(\frac{l j \pi}{n}\right). \quad (4)$$

Hence it suffices to study the action of the coarsegrid scheme on the eigenvectors  $w_l^h$ . It is common for the analysis to distinguish between smooth eigenmodes  $w_l^h$  with  $1 \leq l < n/2$  and the corresponding oscillatory eigenmodes  $w_{l'}^h$ , where  $l' = n - l$ . Defining the constants  $c_l = \cos^2(\frac{l\pi}{2n})$  and  $s_l = \sin^2(\frac{l\pi}{2n})$ , the following general expressions hold

$$TG w_l^h = w_l^h - c_l \frac{\lambda_l^h}{\lambda_{2h}^{2h}} (c_l w_l^h - s_l w_{l'}^h), \quad (5)$$

$$TG w_{l'}^h = w_{l'}^h - s_l \frac{\lambda_{l'}^h}{\lambda_l^{2h}} (c_l w_l^h - s_l w_{l'}^h), \quad (6)$$

where  $\lambda_l^h$  and  $\lambda_{l'}^h$  are the eigenvalues corresponding to  $w_l^h$  and  $w_{l'}^h$  respectively. The famous two-grid analysis of Ernst, Elman and O'leary [14] follows as a smooth mode limit case from expression (5), as for  $l \ll n/2$  we have  $c_l \approx 1$  and  $s_l \approx 0$  implying

$$TG w_l^h \approx \left(1 - \frac{\lambda_l^h}{\lambda_{2h}^{2h}}\right) w_l^h. \quad (7)$$

For notational simplicity we now denote  $\lambda_l^h = \lambda^h$  and  $\lambda_{l'}^h = \lambda^H$ . For the indefinite Helmholtz equation it is well-known that  $\lambda^h = \lambda_L^h - k^2$ , where  $\lambda_L^h \in [0, 2^{d+1}/h^2]$  is the corresponding eigenvalue of the discretised  $d$ -dimensional Laplace operator  $L = -\Delta^h$ . As discussed in [14] a quasi-zero denominator in expression (7) due to  $\lambda_L^h \approx k^2$  and/or oppositely signed eigenvalues  $\lambda^h$  and  $\lambda^H$  can lead to two-grid instability, i.e.

$$\left|1 - \frac{\lambda^h}{\lambda^H}\right| = \left|1 - \frac{\lambda_L^h - k^2}{\lambda_L^H - k^2}\right| \gg 1. \quad (8)$$

Indeed, this observation acted as the main motivation for the development of the CSL and CSG preconditioners, see further. Note however that the eigenvalues  $\lambda_L^h$  and  $\lambda_L^H$  corresponding to the

<sup>†</sup>The standard intergrid operators used in this text are linear interpolation and full weighting restriction.

smoothest eigenmodes  $w_l^h$  resp.  $w_l^{2h}$  with  $l \ll n/2$  are relatively close to zero in comparison to  $k^2$ . Indeed, we have that both  $\lambda_L^h \ll k^2$  and  $\lambda_L^H \ll k^2$ , as  $l \ll n/2$  implies that  $s_l \approx 0$  and  $c_l \approx 1$  such that

$$\lambda_L^h = \frac{4d}{h^2} \sin^2 \left( \frac{l\pi}{2n} \right) = \frac{4d}{h^2} s_l \approx 0 \ll k^2, \quad (9)$$

$$\lambda_L^H = \frac{d}{h^2} \sin^2 \left( \frac{l\pi}{n} \right) = \frac{4d}{h^2} \sin^2 \left( \frac{l\pi}{2n} \right) \cos^2 \left( \frac{l\pi}{2n} \right) = \frac{4d}{h^2} s_l c_l \approx 0 \ll k^2. \quad (10)$$

Hence the smoothest eigenmodes are always projected onto zero by the  $TG$  operator, as (7) implies that

$$TG w_l^h \approx \left( 1 - \frac{\lambda_L^h - k^2}{\lambda_L^H - k^2} \right) w_l^h = \left( 1 - \frac{\lambda_L^h/k^2 - 1}{\lambda_L^H/k^2 - 1} \right) w_l^h \approx 0 \cdot w_l^h. \quad (11)$$

Note that this projection of the smoothest modes onto zero is a desirable property for any coarsegrid correction scheme, as it guarantees optimal collaboration between correction scheme and smoother.

## 2.2. Complex shifted Laplacian

In the renowned 2004 paper [17] Erlangga, Vuik and Oosterlee introduced a solution to the occurrence of small-valued coarsegrid eigenvalues which destroy two-grid stability by means of introducing a real and/or complex shift to the problem. The slightly perturbed Helmholtz problem

$$(-\Delta^h - (\alpha + \beta i)k^2)u^h = \chi^h \quad (12)$$

proposed within [17] can be solved efficiently using multigrid given a sufficiently large complex shift part  $\beta$ . A recent discussion on the exact value of the shift parameter  $\beta$  can be found in [13]. Note that alternatively, the above expression can be reformulated in a slightly less convenient form as  $(-\Delta^h - re^{i\theta}k^2)u^h = \chi^h$  such that the eigenvalues of the perturbed discretisation matrix  $\tilde{A}^h$  are given by  $\lambda^h = \lambda_L^h - re^{i\theta}k^2$ . The two-grid correction scheme on this perturbed problem is

$$TG = (I - I_{2h}^h (\tilde{A}^{2h})^{-1} I_h^{2h} \tilde{A}^h), \quad (13)$$

where  $\tilde{A}^h = L - re^{i\theta}k^2$ . Note that the eigenvectors  $w_l^h$  of the original Helmholtz problem (1) are preserved by shifting the Laplace operator. Hence, following expression (7) for the smooth modes, the denominator of the fraction

$$\left| 1 - \frac{\lambda_L^h - re^{i\theta}k^2}{\lambda_L^H - re^{i\theta}k^2} \right| \quad (14)$$

will never approach zero as  $|\lambda^H| = |\lambda_L^H - re^{i\theta}k^2| > |\mathcal{I}(re^{i\theta}k^2)| > 0$  for all coarsegrid eigenvalues  $\lambda_L^H$  due to the complex part. Similarly to the unperturbed problem, the smoothest eigenvalues are projected onto zero by the  $TG$  operator, as for  $l \ll n/2$  we have  $\lambda_L^h \ll k^2$  and  $\lambda_L^H \ll k^2$  such that

$$TG w_l^h = \left( 1 - \frac{\lambda_L^h - re^{i\theta}k^2}{\lambda_L^H - re^{i\theta}k^2} \right) w_l^h = \left( 1 - \frac{\lambda_L^h/k^2 - re^{i\theta}}{\lambda_L^H/k^2 - re^{i\theta}} \right) w_l^h \approx 0 \cdot w_l^h. \quad (15)$$

Thus for a sufficiently large complex shift, the two-grid correction scheme is stable for the shifted problem (14) and projects the smooth eigenvalues near the origin as requested (15). However, despite its ability to efficiently solve the perturbed problem (12), the original Helmholtz problem (1) remains unsolvable by multigrid, and the multigrid solution method for the CSL problem (12) can only be used as a preconditioner solver to the original Helmholtz problem. The remaining preconditioned matrix-vector system is most commonly solved by means of a Krylov subspace method like e.g. restarted GMRES.

### 2.3. Complex stretched grid

Introduced as an alternative to Complex Shifted Laplacian in 2010 [27], the Complex Stretched Grid (CSG) preconditioner can be seen as an equivalent method of perturbing the Helmholtz problem to resolve the two-grid correction instability issue. As opposed to CSL however, the existing shift  $-k^2$  is left unchanged, whereas the discretisation grid is rotated into the complex plane

$$(-\Delta^h e^{-i\theta} - k^2)u^h = \chi^h. \quad (16)$$

Note that we have deliberately chosen  $r = 1$ , inducing a pure rotation of the grid around  $-k^2 \in \mathbb{R}$  (see Figure 1), i.e. such that the distance between two gridpoints is left unchanged  $|\tilde{h}| = |he^{-i\frac{\theta}{2}}| = |h|$ . To stress the similarity between CSL and CSG, note that the above CSG problem formulation (16) is equivalent to the CSL problem

$$(-\Delta^h - k^2 e^{i\theta})u^h = \chi^h e^{i\theta}. \quad (17)$$

As discussed in [27] this equivalence between CSL and CSG generally holds. The CSL wavenumber perturbation factor  $1 + \beta i$  (which induces no additional real shift and is therefore commonly used in Helmholtz literature, see e.g. [18]) can always be transferred to a CSG setting by choosing  $r = 1/\cos \theta$  and vice versa. As a general remark, note that for small rotation angles  $\theta \approx 0$ , this implies  $r \approx 1$ . In the following it is assumed without loss of generality that  $r \equiv 1$ .

The eigenvalues of the resulting perturbed CSG problem are given by  $\lambda^h = \lambda_L^h e^{-i\theta} - k^2$ . Referring once more to (7), it follows from an argument similar to (14) that the denominator in

$$\left| 1 - \frac{\lambda_L^h e^{-i\theta} - k^2}{\lambda_L^H e^{-i\theta} - k^2} \right| \quad (18)$$

cannot approach zero when the rotation angle  $\theta$  is sufficiently large, hence solving the two-grid instability issue. Additionally, the smooth mode eigenvalues  $\lambda_l^h$  with  $l \ll n/2$  are projected onto zero by the  $TG$  operator as can be readily derived from (7) in analogy to (15). This implies that the perturbed CSG problem can also be solved very efficiently using multigrid. Analogously to the CSL method however, the Complex Stretched Grid approach can *only* be used as a preconditioner to a general Krylov method, as the original Helmholtz problem cannot be solved by solely using multigrid.

## 3. LEVEL-DEPENDENT COARSEGRID CORRECTION SCHEME

In this section we introduce a level-dependent correction scheme, based on the perturbation idea introduced by the CSL and CSG preconditioners. As opposed to the latter methods however, the problem is now gradually perturbed throughout the grid hierarchy, ensuring both stability of the two-grid scheme on all levels as well as solution of the original problem on the finest grid. This leads to an adapted multigrid correction scheme that is *directly* applicable as a solver for the original Helmholtz problem (1). Additionally, it is shown that the error introduced by perturbing the problem does not undermine the functionality of the multigrid scheme.

### 3.1. Definition and analysis

We introduce the notion of a level-dependent correction scheme based on the CSL and CSG schemes. Assuming the total number of levels in the multigrid hierarchy equals  $p$ , the two-grid scheme on the  $m$ -th level in the level-dependent multigrid hierarchy is defined as

$$TG^m = (I - I_{2h}^h (\tilde{A}^{2h})^{-1} I_h^{2h} \tilde{A}^h), \quad 1 \leq m < p, \quad (19)$$

where for the scheme based on CSL perturbation we state that

$$\begin{aligned} \tilde{A}^h &= -\Delta^h - k^2 e^{i\theta_{m-1}}, \\ \tilde{A}^{2h} &= -\Delta^{2h} - k^2 e^{i\theta_m}, \quad 1 \leq m < p, \end{aligned} \quad (20)$$

or equivalently for the CSG-based variant

$$\begin{aligned}\tilde{A}^h &= -\Delta^h e^{-i\theta_{m-1}} - k^2, \\ \tilde{A}^{2h} &= -\Delta^{2h} e^{-i\theta_m} - k^2, \quad 1 \leq m < p.\end{aligned}\tag{21}$$

The level-dependent parameter  $\theta_m$  is most commonly defined as  $\theta_m = md\theta$  for a *small* perturbation parameter  $d\theta \in [0, \pi]$ , however other perturbations schemes are possible. In this work the per-level perturbation  $d\theta$  is defined as  $d\theta = \theta_{max}/p$  where the fixed angle  $\theta_{max}$  is the maximal rotation angle of the coarsest level, usually chosen  $\theta_{max} = \pi/6$ . Note that, by definition, on the finest level (where  $m = 1$ ) we have  $\tilde{A}^h = A^h = -\Delta^h - k^2$ . This observation acts as our primary motivation for the adaptation of the two-grid scheme into a level-dependent shift or rotation angle rather than a fixed perturbation of the problem on all levels. The hierarchy of two-grid schemes proposed above is designed to ultimately solve the *original* Helmholtz problem on the finest grid.

On every grid the current solution  $v^h$  is corrected by  $v^h \leftarrow v^h + \tilde{e}^h$  using the interpolated solution  $\tilde{e}^h$  of a slightly perturbed coarsegrid problem as an approximation to the exact error. Note that even in a regular two-grid scheme, the transfer of the residual to its discrete coarsegrid representation results in an *approximation* of the finegrid problem (due to restriction and interpolation). In the level-dependent scheme, we minorly alter the problem definition on the coarse grid by adding a *small* shift or domain rotation, obtaining a slightly different approximation to the error. In doing so, we introduce a small additional error on the correction due to the difference in problem definition between coarse and fine grid. In the next paragraphs however it will be shown that the resulting two-grid scheme does not excite smooth modes in the correction term. Indeed, we will prove that the additional error component introduced mainly consists of oscillatory eigenmodes, which are subsequently damped by the smoother, yielding an accurate overall error correction scheme.

*3.1.1. Two-grid analysis of the level-dependent scheme.* In the following paragraph we perform a standard two-grid analysis of the new level-dependent scheme, cfr. Section 2. We distinguish between the CSL- and CSG-based level-dependent schemes as presented above. Focussing first on the CSL scheme and using expression (7), we observe that the denominator of the fraction

$$\left| 1 - \frac{\lambda_L^h - e^{i\theta_{m-1}} k^2}{\lambda_L^H - e^{i\theta_m} k^2} \right| \tag{22}$$

never approaches zero for  $\theta_m$  sufficiently large, resolving the instability issues of the standard two-grid scheme. Additionally, the amplification factor for the smoothest eigenmodes  $w_l^h$  with  $l \ll n/2$  can be calculated as follows

$$\begin{aligned}TG^m w_l^h &= \left( 1 - \frac{\lambda_L^h - e^{i\theta_{m-1}} k^2}{\lambda_L^H - e^{i\theta_m} k^2} \right) w_l^h = \left( 1 - \frac{\lambda_L^h/k^2 - e^{i\theta_{m-1}}}{\lambda_L^H/k^2 - e^{i\theta_m}} \right) w_l^h \\ &\approx (1 - e^{-id\theta}) w_l^h,\end{aligned}\tag{23}$$

where we again assume that both  $\lambda_L^h \ll k^2$  and  $\lambda_L^H \ll k^2$  for the smoothest eigenmodes. Notably, the amplification factor of the smoothest eigenmodes depends on the parameter  $d\theta$ , which in general is very small, implying  $(1 - e^{-id\theta}) \approx 0$  such that the smoothest eigenvalue is indeed projected closely to zero. From (23) it can be derived that stability for the smooth eigenmodes is guaranteed as long as  $d\theta < \pi/3$ , which in practice will always be satisfied. Contrary to the standard two-grid schemes discussed in Section 2 however, the smoothest modes are not *exactly* projected onto zero by the CSL-based  $TG^m$  operator. Instead, the minimal distance between the origin and a smooth mode's projection under the CSL variant of  $TG^m$  is  $|1 - e^{-id\theta}|$ .

Subsequently turning to the level-dependent CSG scheme, it is clear that this scheme is also generally stable as the denominator of

$$\left| 1 - \frac{\lambda_L^h e^{-i\theta_{m-1}} - k^2}{\lambda_L^H e^{-i\theta_m} - k^2} \right| \tag{24}$$



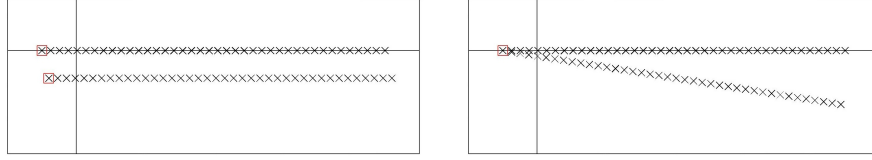


Figure 1. Schematic representation of the spectra of the Dirichlet bounded discretisation matrix operator  $A^h$  and the perturbed CSL (left) and CSG (right) operators  $\tilde{A}^h$ . Note how the CSG perturbation does not change the location of the leftmost (smoothest) eigenvalue, whereas CSL translates the entire spectrum.

does not approach zero for  $\theta_m > 0$  sufficiently large, due to the complex part. Assuming once more that the eigenvalues corresponding to the smoothest eigenmodes  $w_l^h$  (with  $l \ll n/2$ ) are relatively close to zero in comparison to  $k^2$ , such that  $\lambda_L^h \ll k^2$  and  $\lambda_L^H \ll k^2$ , the action of  $TG^m$  on a very smooth eigenmode is given by

$$TG^m w_l^h = \left(1 - \frac{\lambda_L^h e^{-i\theta_{m-1}} - k^2}{\lambda_L^H e^{-i\theta_m} - k^2}\right) w_l^h = \left(1 - \frac{(\lambda_L^h/k^2)e^{-i\theta_{m-1}} - 1}{(\lambda_L^H/k^2)e^{-i\theta_m} - 1}\right) w_l^h \approx 0 \cdot w_l^h. \quad (25)$$

Consequently, one observes that using the CSG scheme the smoothest eigenvalues are projected onto zero as requested, independently of the parameter  $d\theta$ . Note that this was not the case for the CSL variant of  $TG^m$ , see (23). This subtle difference between the CSL and CSG level-dependent schemes can be clarified further by studying their fundamental operation on the spectrum of the discretisation matrix  $A^h$ , see Figure 1: while CSL induces a shift of the entire finegrid spectrum over a distance defined by  $d\theta$ , causing the fine- and coarsegrid eigenvalues in the level-dependent scheme to be distinctly different, CSG rotates the spectrum over an angle  $-d\theta$  around the point  $-k^2 \in \mathbb{R}$ , such that the smoothest corresponding eigenvalues on the fine and coarse grid are nearly left unchanged. This signifies that for smooth eigenmodes the CSG level-dependent scheme (21) resembles the action of the standard coarsegrid correction scheme (2) somewhat closer than the CSL variant. In light of this observation, our preference in solving practical problem goes out to the CSG variant of the level-dependent scheme.

**3.1.2. Spectral analysis of the error components.** A similar spectral analysis can be performed from a slightly different point of view. In the following we focus without loss of generality on the coarsegrid residual equation of the  $TG^1$  correction scheme

$$\tilde{A}^{2h} \tilde{e}^{2h} = \tilde{r}^{2h}, \quad (26)$$

where the righthand side  $\tilde{r}^{2h} = c r^{2h} = c I_h^{2h} r^h$  ( $c \in \mathbb{C}$ ) is the (scaled) restricted finegrid residual  $r^h = f^h - A^h v^h$ . The scaling of the righthand side is mandatory for the CSG variant of the scheme, and can intuitively be seen as a natural effect to the rotation of the spectrum, i.e. typically  $c = e^{-id\theta}$ , see also (17). For the CSL variant, no scaling is required as the spectrum is merely translated in the complex plane, thus setting  $c = 1$ . Equation (26) is solved for  $\tilde{e}^{2h}$  which is then interpolated to correct the finegrid solution. Comparing to the standard two-grid correction scheme with residual equation

$$A^{2h} e^{2h} = r^{2h}, \quad (27)$$

the solution  $\tilde{e}^{2h}$  to the perturbed residual equation (26) differs slightly from  $e^{2h}$  due to the minor difference in the level-dependent scheme between the fine- and coarsegrid operators. The relation between  $\tilde{e}^{2h}$  and  $e^{2h}$  is given by equating the above, i.e.  $\tilde{A}^{2h} \tilde{e}^{2h} = c A^{2h} e^{2h}$ , which implies

$$\begin{aligned} \tilde{e}^{2h} &= c \tilde{A}^{2h-1} A^{2h} e^{2h} = c \tilde{A}^{2h-1} (A^{2h} + c^{-1} \tilde{A}^{2h} - c^{-1} \tilde{A}^{2h}) e^{2h} \\ &= e^{2h} + \left( \tilde{A}^{2h-1} (c A^{2h} - \tilde{A}^{2h}) \right) e^{2h}. \end{aligned} \quad (28)$$

As briefly pointed out in the introduction to this section, the level-dependent scheme indeed introduces a small additional error  $\varepsilon^{2h}$  on the correction due to the difference in problem definition between coarse and fine grid. This error can effectively be characterised using equation (28) as

$$\varepsilon^{2h} = \left( \tilde{A}^{2h-1} (c A^{2h} - \tilde{A}^{2h}) \right) e^{2h}, \quad (29)$$

and preferably equals zero. In reality however,  $e^{2h}$  consists of a linear combination of coarsegrid eigenmodes, which implies that the additional error  $\varepsilon^{2h}$  is given by

$$e^{2h} = \sum_{l=1}^{n/2} a_l w_l^{2h} \quad \Rightarrow \quad \varepsilon^{2h} = \sum_{l=1}^{n/2} \gamma_l^{2h} a_l w_l^{2h}, \quad (30)$$

for coefficients  $a_l \in \mathbb{R}$ . The weights  $\gamma_l^{2h}$  are the eigenvalues of the matrix operator  $\tilde{A}^{2h-1} (c A^{2h} - \tilde{A}^{2h})$ , given explicitly by

$$\gamma_l^{2h} = \frac{c \lambda_l^{A^{2h}} - \lambda_l^{\tilde{A}^{2h}}}{\lambda_l^{\tilde{A}^{2h}}}, \quad l = 1, \dots, \frac{n}{2}. \quad (31)$$

These eigenvalues can be seen as weighting each eigenmode component of the additional error  $\varepsilon^{2h}$  implied by the perturbation of the problem. They are shown in Figure 2 for various problem settings.

For the CSL level-dependent scheme without additional righthand-side scaling ( $c = 1$ ), the eigenvalues  $\gamma_l^{2h}$  can be rewritten explicitly as

$$\gamma_l^{2h} = \frac{\lambda_l^L - k^2 - \lambda_l^L + k^2 e^{id\theta}}{\lambda_l^L - k^2 e^{id\theta}} = \frac{k^2 (e^{id\theta} - 1)}{\lambda_l^L - k^2 e^{id\theta}}, \quad (32)$$

where the symbol  $L$  is now used to denote the discretised *coarsegrid* Laplacian  $L = -\Delta^{2h}$ . Note that for the Poisson problem with  $k^2 = 0$ , the above expression reduces to

$$\gamma_l^{2h} = 0, \quad l = 1, \dots, \frac{n}{2}, \quad (33)$$

hence considering (30) we have  $\varepsilon^{2h} = 0$ . This result implies that for a Poisson problem the level-dependent correction scheme is theoretically identical to the original multigrid correction scheme, as no additional error is added by solving the equation on a perturbed coarser level.

Moving back to the general Helmholtz setting where  $k^2$  is distinctly different from zero, we consider a very smooth coarsegrid eigenmode  $w_l^{2h}$  with  $l \ll n/4$  which implies  $\lambda_l^L \ll k^2$ . For such a smooth eigenmode it follows from (32) that

$$\gamma_l^{2h} \approx e^{-id\theta} - 1, \quad l \ll \frac{n}{4}. \quad (34)$$

Considering (30) and presuming that the per-level perturbation  $d\theta$  is small, this result implies that smooth eigenmodes only marginally contribute to the error component  $\varepsilon^{2h}$ . Consequently, the additional error component  $\varepsilon^{2h}$  that was induced by the adaptation of the standard correction scheme to a level-dependent scheme consists mainly of *oscillatory* eigenmodes. As the elimination of the oscillatory eigenmodes is exactly the function of the pre- and postsmoothing steps encapsulating the correction scheme<sup>‡</sup>, the entire smoothed two-grid scheme is expected to perform (nearly) as efficient as the standard two-grid schemes currently used as preconditioners. We again stress however that the newly developed level-dependent scheme can be used *directly* as a solver for the indefinite problem, instead of suiting preconditioning purposes.

<sup>‡</sup>Moreover, the coefficients  $a_l$  for oscillatory modes  $w_l^{2h}$  in expression (30) can in fact a priori be assumed small due to presmoothing.



We additionally note that, given a sufficiently fine discretisation with respect to the wavenumber  $k$  (e.g. respecting the  $kh < 0.625$  criterion, see [5]), the spectral radius of the eigenvalues  $\gamma_l^{2h}$  is bounded as a function of the number of gridpoints  $n$ . Given a fixed wavenumber  $k$  and perturbation angle  $d\theta$ , the spectral radius is maximal for the eigenvalue  $\gamma_l^{2h}$  minimizing the denominator  $|\lambda_l^L - k^2 e^{id\theta}|$  of (32). This minimal distance over all  $l$  however remains quasi unchanged for increasingly fine meshes, as the spectrum of the Laplacian  $L$  only expands along the positive real line as  $n$  grows. As a consequence of this observation, good scalability in function of  $n$  can be expected for the level-dependent scheme.

A similar analysis can be performed for the CSG level-dependent scheme, yet to obtain comparable properties to those following from expression (32), the righthand side  $r^{2h}$  should be scaled by  $c = e^{-id\theta}$ , i.e. the same rotation used to perturb the matrix operator  $\Delta^{2h}$ . Note that this scaling is quite natural, see (17), and it does not complicate the coarsegrid system; however it will make sure the Poisson problem is solved identically by the standard- and CSG level-dependent schemes, cfr. (33).

For the CSG level-dependent scheme with righthand-side scaling parameter  $c = e^{-id\theta}$ , the eigenvalues  $\gamma_l^{2h}$  (31) can be rewritten as

$$\gamma_l^{2h} = \frac{\lambda_l^L e^{-id\theta} - k^2 e^{-id\theta} - \lambda_l^L e^{-id\theta} + k^2}{\lambda_l^L e^{-id\theta} - k^2} = \frac{k^2(1 - e^{-id\theta})}{\lambda_l^L e^{-id\theta} - k^2}, \quad (35)$$

which is identical to expression (32). Consequently the properties stated above hold: the CSG level-dependent scheme with  $c = e^{-id\theta}$  is identical to the standard correction scheme when solving the Poisson problem, and the additional error  $\varepsilon^{2h}$  that rises when applying the level-dependent scheme to a general Helmholtz problem consists mainly of oscillatory modes, which are consecutively eliminated by the application of the smoother.

The eigenvalues  $\gamma_l^{2h}$  characterising the error  $\varepsilon^{2h}$  between the original and the level-dependent multigrid schemes are plotted in Figure 2. The smooth mode limit value  $(e^{-id\theta} - 1)$  from (34) is marked by the  $\square$ -symbol. The eigenvalue corresponding to the smoothest eigenmode is indeed found closely to zero as suggested by the discussion above. The oscillatory eigenmodes clearly have the largest contribution in the perturbation error  $\varepsilon^{2h}$ . Note that the spectrum does not expand in function of the number of discretisation points  $n$ . The numerical experiments covered in section 4 indeed confirm that scalability in function of the number of gridpoints is maintained. As shown by the figures, the spectrum does however expand in function of the wavenumber  $k$ , implying perfect  $k$ -scalability cannot be assured for the level-dependent scheme.<sup>§</sup> The effects caused by the growing spectral radius in function of  $k$  are however somewhat extenuated by the following two observations. First, the growing addition of oscillatory modes by the correction scheme is generally not problematic, as they are efficiently eliminated by the smoother, and second, the per-level perturbation  $d\theta$  decreases in function of the number of grid levels, such that the contribution of the smooth modes to  $\varepsilon^{2h}$  (which *would* be problematic if growing) diminishes in function of  $n$ . As a final remark we note from Figure 2 that the spectrum is much more clustered around the origin when using non-Dirichlet boundary conditions like e.g. ECS, Sommerfeld or PML boundaries, which is particularly important regarding the contribution of smooth modes to  $\varepsilon^{2h}$ . Indeed, experiments have shown that the level-dependent scheme performs very poorly on a purely Dirichlet-bounded problem. As often suggested in the literature, the use of Dirichlet boundary conditions implies a ‘worst case scenario’ for efficient solution, as no natural damping occurs. Hence, it must be noted that the level-dependent scheme is mainly devised for use on problems with more realistic ECS or Sommerfeld boundary conditions.

<sup>§</sup>At the moment of writing the authors are unaware of any iterative method for the Helmholtz problem that ensures perfect  $k$ -scalability. However, significant efforts towards this aim for the class of preconditioned Krylov solvers were recently made by Vuik et al., see [32].

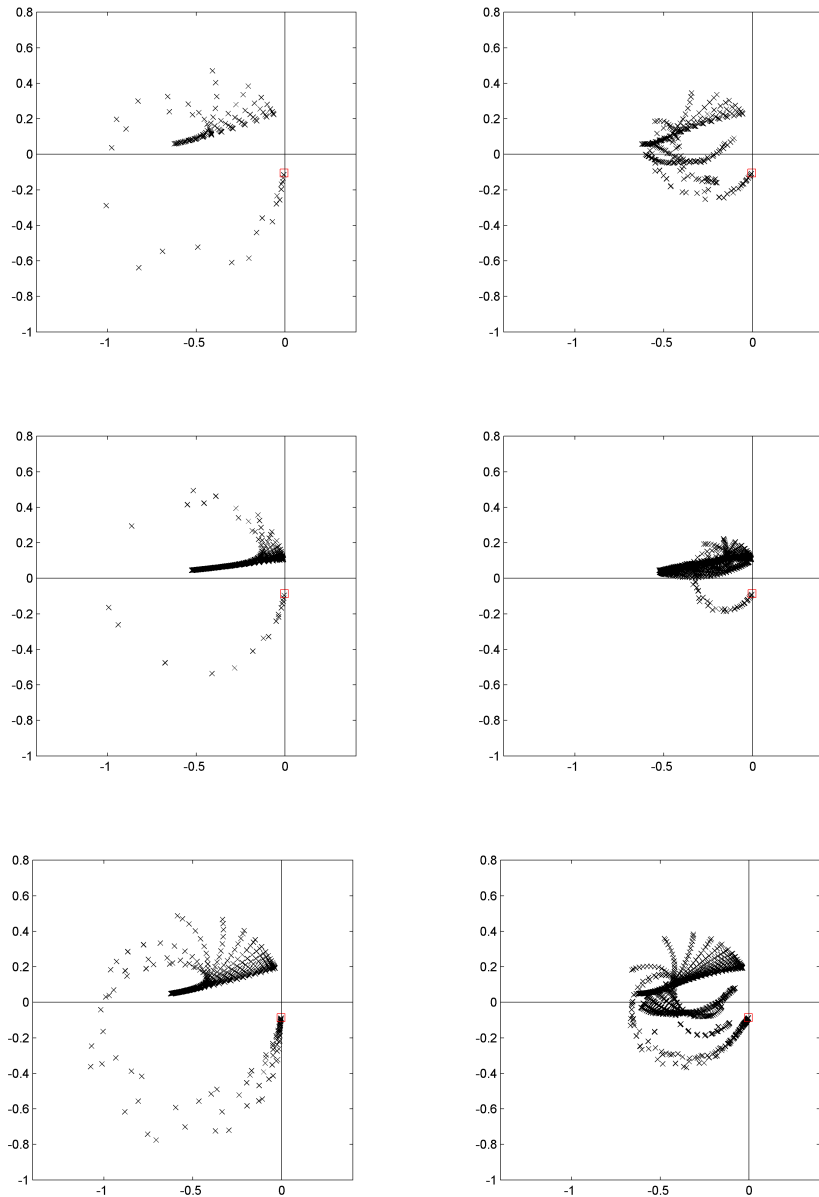


Figure 2. Spectrum of the eigenvalues  $\gamma_l^{2h}$  (31) for a 2D constant- $k$  Helmholtz problem with Dirichlet (left) and ECS (right) boundary conditions. Top: wavenumber  $k = 20$  and  $n = 32$  discretisation points. Mid: wavenumber  $k = 20$  and  $n = 64$  discretisation points. Bottom: wavenumber  $k = 40$  and  $n = 64$  discretisation points.

### 3.2. On the relaxation method or ‘smoother’

As the second component of the multigrid scheme, the smoother plays a crucial role in the elimination of highly oscillatory modes in the error. The smoothing of oscillatory eigenvalues is fundamental for the performance of the multigrid scheme, as it complements the action of the coarsegrid correction scheme constructed above.

**3.2.1. Introduction of polynomial smoothers.** Being a basic iterative scheme, the smoother can be chosen from a wide variety of standard relaxation methods, the most common being weighted Jacobi relaxation given by the iteration matrix

$$R_\omega = I - \omega D^{h^{-1}} A^h, \quad (36)$$

where  $\omega \in \mathbb{C}$ . Note that in case of a constant diagonal  $D^h$ , the above expression can be rewritten as the evaluation of a first order polynomial  $p_1(t)$  in  $A$ , where

$$p_1(t) = I - \omega_1 t, \quad (37)$$

with  $\omega_1 \in \mathbb{C}$ . This observation led to the development of so-called *polynomial smoothers*, introduced in [37] and first embedded in multigrid in e.g. [3], in which a higher-order polynomial  $p_m(t)$  of degree  $m \geq 1$  is employed as a smoother

$$p_m(t) = (I - \omega_1 t)(I - \omega_2 t) \dots (I - \omega_m t) = I + \sum_{i=1}^m c_i t^i, \quad (38)$$

where  $\omega_1, \dots, \omega_m \in \mathbb{C}$  and  $c_1, \dots, c_m \in \mathbb{C}$ . Note that per definition  $p_m(0) = 1$ . As described in [28], the following two features are essential properties of any smoothing polynomial  $p_m(t)$ :

- (1) Stability property:  $\forall h, \forall \lambda_l \in \text{spec}(A^h) : |p_m(\lambda_l)| < 1 \quad (l = 1, \dots, n)$ ,
- (2) Smoothing property:  $p_m(\lambda_n) = 0$ .

The first condition expresses the smoother  $p_m(t)$  is preferably stable on *all* levels, meaning none of the eigenmodes are amplified by its application. The second condition implies the smoother  $p_m(t)$  effectively damps the oscillatory eigenmodes, projecting the corresponding high-frequency eigenvalues onto (or close to) zero.

In view of these conditions we suggest GMRES(m) as a replacement for the standard smoothing schemes, where the degree  $m$  is typically chosen  $m = 3$ . Indeed, this implies the intrinsic construction of a smoothing polynomial of degree  $m$  for which the coefficients  $c_1, \dots, c_m$  are optimally chosen with respect to a minimisation of the current residual. Note that the use of GMRES(m) implies a *level-dependent smoother* is included in the multigrid scheme, as the polynomial coefficients are redetermined upon every application of the smoother.

**3.2.2. Motivation for the use of GMRES(m).** Although the application of a polynomial smoother  $p_m(t)$  is computationally more expensive than standard weighted Jacobi or Gauss-Seidel relaxation, the use of an  $m$ -th order polynomial with variable polynomial coefficients yields significant advantages over the fixed first-order Jacobi smoothing polynomial (37). Indeed, consider the  $\omega$ -Jacobi smoother with fixed relaxation parameter  $\omega \in \mathbb{C}$  as given by (36). Stability of the smoother is guaranteed whenever

$$|1 - \omega \lambda_l^h| < 1, \quad \forall l = 1, \dots, n, \quad \forall h, \quad (39)$$

on all levels in the multigrid hierarchy. Note that on the finest grid ( $h$ -level) the eigenvalues  $\lambda_l^h$  corresponding to the oscillatory eigenmodes  $w_l^h$  ( $l > n/2$ ) are distinctively situated on the positive real line<sup>¶</sup>, i.e.  $\lambda_l^h \gg 0$ , implying one at least requires  $\mathcal{R}(\omega) > 0$  in order to obtain a stable smoother. However, when the wavenumber  $k$  is significantly larger than zero, there notably exists a coarse  $H$ -level in the grid hierarchy on which all eigenvalues  $\lambda_l^H$  are negative, meaning  $\forall l : \lambda_l^H \ll 0$ . On this level one clearly requires  $\mathcal{R}(\omega) < 0$  for (39) to hold, resulting in a contradictory overall requirement on the relaxation parameter  $\omega$ . As a direct generalisation of this result one can state that fixing the coefficients of a polynomial smoother inevitably results in smoother instability on certain levels in

<sup>¶</sup>To avoid technicalities we still assume that the problem under consideration has Dirichlet bounds, implying all eigenvalues of  $A^h$  are situated on the real line.

the multigrid hierarchy. This stability issue is directly resolved by using GMRES(m), which selects different coefficients on each level.

In addition to the stability issues of a polynomial smoother with fixed coefficients, one can furthermore show that for a given level in the multigrid hierarchy, a first (and even second) order polynomial can never lead to a stable smoother. Due to the indefinite nature of the problem the smoothest eigenvalues  $\lambda_l^h$  ( $l \ll n/2$ ) on the finest levels are located on the negative real line, i.e.  $\lambda_l^h \ll 0$ , whereas for the oscillatory eigenvalues  $\lambda_{l'}^h$  ( $l' > n/2$ ) one has  $\lambda_{l'}^h \gg 0$ . Note that for the smoother to be stable, it follows from (39) that  $\forall l : \mathcal{R}(\omega)\lambda_l^h \in ]0, 2[$ . For positive eigenvalues  $\lambda_l^h > 0$  this implies that  $0 < \mathcal{R}(\omega) < 2/\lambda_l^h$ , whereas for  $\lambda_l^h < 0$  one requires  $2/\lambda_l^h < \mathcal{R}(\omega) < 0$ , which are clearly contradictory requirements. Consequently, no choice for  $\omega \in \mathbb{C}$  can possibly lead to a stable first order polynomial smoother.

One of the first papers to partially substitute the smoother by GMRES(m) is notably [14]. In [12] a GMRES(m) method was used as a smoother for the entire multigrid cycle. In [28] it was shown that for a Helmholtz problem with ECS boundaries a stable particular polynomial  $p_m(t)$  of degree  $m = 3$  which effectively maps the most oscillatory eigenvalues onto zero can always be constructed. Since the experiments conducted further on in this paper apply ECS boundaries to represent outgoing wave boundary conditions, we have chosen to use GMRES(3) as a smoother substitute throughout this work. Note that when using GMRES(m) as a smoother substitute within a preconditioning multigrid method, the multigrid cycle intrinsically is a variable nonlinear preconditioner which must be combined with a *flexible* outer Krylov subspace method, see [33].

#### 4. NUMERICAL EXPERIMENTS

In this section we extensively test the convergence and scalability of our new level-dependent multigrid scheme with respect to the number of finegrid discretisation points  $n$ , the wavenumber  $k$  and the per-level perturbation parameter  $d\theta$ . Additionally, we illustrate the convergence speed of the new multigrid method and compare with current state-of-the-art solution methods like e.g. Complex Shifted/Stretched preconditioned GMRES in terms of iteration count and CPU time. The three benchmark problems used cover a wide variety of possible Helmholtz settings, providing an adequate testing framework for the new level-dependent multigrid solution method. All experiments are conducted in a 2D setting and use ECS absorbing boundary conditions, unless explicitly stated otherwise.<sup>||</sup>

##### 4.1. The constant $k$ model

The first model problem is the most elementary Helmholtz equation and is therefore a grateful subject for theoretical analysis and a benchmark for numerical experiments. It describes a general scattering problem in a homogeneous medium, i.e. with a constant wave number  $k$  and a point-source located at the center of the domain  $\Omega$ . The solution vanishes towards infinity in all directions. We use the following standard setting of parameters:

$$\begin{aligned} \Omega &= (0, 1)^2, \\ k(x, y) &= k^2, \\ \chi(x, y) &= \begin{cases} 1, & \text{for } x = y = 1/2, \\ 0, & \text{elsewhere,} \end{cases} \\ u &= \text{outgoing on } \partial\Omega. \end{aligned}$$

In a dimensionless formulation, the computational domain  $\Omega$  is a unit square with absorbing boundary conditions on all four edges. The domain is discretised using  $n$  equidistant gridpoints

<sup>||</sup>Machine specifications: Intel® Core™ i7-2720QM 2.20GHz CPU, 6MB Cache, 8GB RAM

$n_x \times n_y$	$32^2$	$64^2$	$128^2$	$256^2$	$512^2$	$1024^2$
$d\theta$	$\frac{\pi}{30}$	$\frac{\pi}{36}$	$\frac{\pi}{42}$	$\frac{\pi}{48}$	$\frac{\pi}{54}$	$\frac{\pi}{60}$
MG iter	77	33	25	25	25	28
CPU total	0.96	1.12	2.81	12.78	53.78	248.8
CPU/1000	0.94	0.27	0.17	0.20	0.21	0.24

Table I. Multigrid performance of the level-dependent scheme for a 2D constant  $k$  Helmholtz problem with  $k = 40$ . Listed are the number of iterations, total CPU time until convergence and CPU time per 1000 gridpoints (in s.) for different levels of discretisation. A series of V(1,1)-cycles with GMRES(3) smoothing is used as a solver.

$n_x \times n_y$	$64^2$	$128^2$	$256^2$	$512^2$	$1024^2$	$2048^2$
$d\theta$	$\frac{\pi}{36}$	$\frac{\pi}{42}$	$\frac{\pi}{48}$	$\frac{\pi}{54}$	$\frac{\pi}{60}$	$\frac{\pi}{66}$
MG iter	180	57	39	40	40	43
CPU total	5.78	6.26	19.99	86.45	357.2	1552
CPU/1000	1.41	0.38	0.31	0.33	0.34	0.37

Table II. Multigrid performance of the level-dependent scheme for a 2D constant  $k$  Helmholtz problem with  $k = 80$ . Listed are the number of iterations, total CPU time until convergence and CPU time per 1000 gridpoints (in s.) for different levels of discretisation. A series of V(1,1)-cycles with GMRES(3) smoothing is used as a solver.

in every spatial dimension, with boundary conditions implemented by an ECS layer consisting of  $n/4$  gridpoints surrounding the domain.

Table I and II show some level-dependent multigrid method convergence results for the constant- $k$  model problem. The number of multigrid V(1,1)-cycles and corresponding CPU time to solve the problem to a residual tolerance of  $10^{-7}$  are given in function of the finegrid discretisation\*\*. As can be read in standard textbooks [11, 35], scalability in function of  $n$  (sometimes referred to as ‘ $h$ -scalability’) is a requirement for any multigrid method, i.e. the number of iterations is expected to remain constant in function of the finelevel discretisation. For the new level-dependent method one indeed observes that scalability in function of  $n$  is guaranteed. Note how the total CPU time scales with the number of finelevel gridpoints  $n$ , as doubling the number of gridpoints implies CPU time  $\times 2^d$ .

Note how convergence deteriorates at very rough finegrid discretisations for which  $kh > 0.625$ , i.e. we do not meet the requirements of 10 gridpoints per wavelength on the finest level of the multigrid hierarchy. Is is well-known (see [5]) that from a physical point of view traditional multigrid methods require a minimum number of gridpoints per wavelength to accurately represent (all eigenmodes of) the solution on the finest level. Additionally, this condition translates into a direct theoretical requirement for the level-dependent scheme due to the definition of the per-level perturbation parameter  $d\theta = \theta_{max}/p$  where  $p$  is the total number of levels in the hierarchy and  $\theta_{max}$  is a fixed rotation angle (commonly chosen  $\theta_{max} = \pi/6$ ). Consequently, as the coarsegrid correction requires  $d\theta$  to be sufficiently small, see (34), convergence of the level-dependent scheme clearly benefits from a fairly fine discretisation. We refer to Table V for a more elaborate discussion on the impact of the perturbation parameter  $d\theta$  on convergence.

Tables III and IV compare performance of the level-dependent multigrid method to the current state-of-the-art CSG preconditioned Krylov methods, which is (to certain extend) equivalent to the CSL-preconditioned Krylov method as proven in [27]. Note that in each outer Krylov step one V(1,1)-cycle is used to approximately solve the preconditioning system, as is common practice in the literature [18, 19, 20, 36]. The LVL-MG iterations counter appoints the required number of V(1,1)-cycles to solve the problem to a residual tolerance of  $10^{-7}$ , whereas the MG-Krylov

\*\*Note that a finegrid number of discretisation points  $n = 2^p$  corresponds to a number of  $p$  levels in the multigrid hierarchy for a full V-cycle.

$k$	20	40	80	160	320
$n_x \times n_y$	$32^2$	$64^2$	$128^2$	$256^2$	$512^2$
MG-FGMRES	19 (0.37)	29 (1.20)	53 (8.09)	106 (125)	204 (1605)
MG-FGMRES(10)	21 (0.38)	30 (1.13)	62 (7.48)	125 (72.9)	249 (625)
LVL-MG	22 (0.37)	33 (1.12)	57 (6.26)	111 (56.7)	224 (488)

Table III. 2D Constant- $k$  problem with ECS boundary conditions. Table comparing iterations and CPU time (in s.) for standard CSG-preconditioned GMRES and level-dependent multigrid. A V(1,1)-cycle with GMRES(3) smoother is used for both preconditioning and level-dependent multigrid. Wavenumber-dependent discretisation following the  $kh < 0.625$  criterion for a minimum of 10 gridpoints per wavelength.

$k$	20	40	80	160	320
$n_x \times n_y$	$32^2$	$64^2$	$128^2$	$256^2$	$512^2$
MG-FGMRES	17 (0.27)	36 (0.88)	73 (5.65)	146 (56.2)	291 (1262)
MG-FGMRES(10)	23 (0.32)	41 (0.88)	77 (4.33)	164 (35.2)	306 (331)
LVL-MG	23 (0.30)	36 (0.73)	64 (3.33)	119 (23.1)	237 (222)

Table IV. 2D Constant- $k$  problem with Sommerfeld radiating boundary conditions. Table comparing iterations and CPU time (in s.) for standard CSG-preconditioned GMRES and level-dependent multigrid. A V(1,1)-cycle with GMRES(3) smoother is used for both preconditioning and level-dependent multigrid. Wavenumber-dependent discretisation following the  $kh < 0.625$  criterion for a minimum of 10 gridpoints per wavelength.

iterations represent the number of outer Krylov iterations. The computational cost of the latter is thus the combined cost of the preconditioning V-cycles and the Krylov steps.

It is clear from the table that the scalability of the LVL-MG method is comparable to that of the MG-FGMRES solver in terms of iteration count. However, as the LVL-MG method is a directly applicable multigrid method, the entire computational cost of the outer Krylov method<sup>††</sup> is dropped when considering the LVL-MG approach, which is clearly reflected in the CPU time. Indeed, LVL-MG significantly outperforms the non-restarted MG-FGMRES solver in terms of computational time for large systems. The occasional increase in iterations (thus in V-cycles) is negligible compared to the computational gains from directly using level-dependent multigrid as a solver. Comparing our level-dependent multigrid solver to the fast restarted MG-GMRES(10) method, one observes an effective speed-up ranging from at least 5% up to over +30% for large-scale problems. Summarizing, it can be stated that the LVL-MG method is remarkably more scalable in terms of CPU time than the current state-of-the-art preconditioned Krylov methods, in particular for large scale problems with high wavenumbers.

Note that for both methods good scalability in function of the wavenumber  $k$  (referred to as ‘ $k$ -scalability’) is de facto not guaranteed. The rising computational cost of MG-GMRES with ECS boundaries in function of  $k$  was explained in [26] by pointing out the convergence rate behaves asymptotically when nearing  $k^2 = 4/h^2$  (independently of the problem dimension). A comparable argument explains the similarly poor  $k$ -scalability for the level-dependent multigrid scheme. Indeed, increasing  $k$  might cause the spectrum on some of the finest (only slightly perturbed) grids to contain eigenvalues which are located closer to zero, as illustrated by Figure 3, resulting in a slight decline in multigrid convergence. However, the level-dependent scheme clearly has improved  $k$ -scalability over the MG-GMRES scheme. A combination of the new level-dependent scheme with e.g. multigrid deflation as proposed in [32] might improve  $k$ -scalability even further.

In Table V the relation between the level-dependent multigrid convergence and the perturbation parameter  $d\theta$  is shown. It is clear that for very small  $d\theta$  the level-dependent scheme performs poorly.

<sup>††</sup>We remark that the total computational cost of a single MG-FGMRES iteration is *strictly greater* than the cost of a LVL-MG solver iteration, as MG-FGMRES requires both an additional matvec plus an orthogonalisation procedure to be executed in comparison to the LVL-MG method. Additionally, a similar conclusion holds for the storage costs, which is due to the storage of the Krylov base vectors, yielding a total of  $m + 1$  times the storage cost of the finegrid solution, whereas a multigrid scheme uses a maximum of two times this storage.



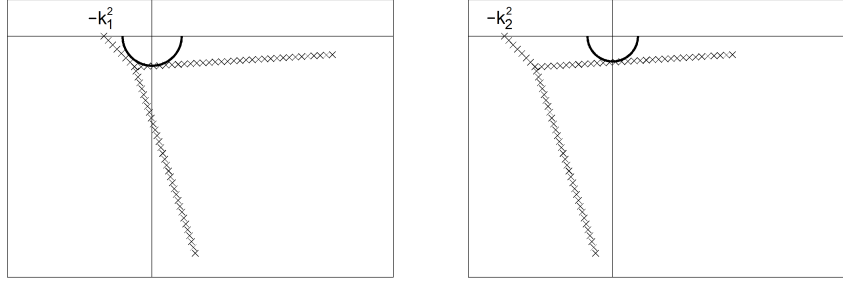


Figure 3. Schematic representation of the spectra of the perturbed ECS bounded CSG discretisation matrix operator  $\tilde{A}^h$  for different values of the wavenumber  $k = k_1$  (left) and  $k = k_2$  (right). Note that  $k_1 < k_2$  implies that the righthand-side spectrum is generally located closer to zero compared to the lefthand-side.

$\theta_{max}$	$\frac{\pi}{15}$	$\frac{\pi}{12}$	$\frac{\pi}{10}$	$\frac{\pi}{8}$	$\frac{\pi}{6}$	$\frac{\pi}{5}$	$\frac{\pi}{4}$
$d\theta$	$\frac{\pi}{105}$	$\frac{\pi}{84}$	$\frac{\pi}{70}$	$\frac{\pi}{56}$	$\frac{\pi}{42}$	$\frac{\pi}{35}$	$\frac{\pi}{28}$
MG iter	27	25	23	22	22	25	28
CPU total	3.05	2.83	2.61	2.50	2.51	2.84	3.15

Table V. Convergence dependency of level-dependent multigrid on the per-level perturbation parameter  $d\theta$ . Shown are the number of iterations and CPU time (in s.) for a 2D constant  $k$  Helmholtz problem with wavenumber  $k = 30$  and a fine-level discretisation with  $n = 128$  gridpoints. A level-dependent V(1,1)-cycle with GMRES(3) smoother is used as a solver.

Indeed, by using a too small rotation or shift some coarselevel spectra might be insufficiently rotated/shifted away from zero, causing the denominator of (22) for the corresponding level(s) yet to approach zero. This possibly leads to a highly instable correction scheme, which is detrimental for the multigrid convergence. A very large per-level perturbation  $d\theta$  on the other hand apparently also gives rise to worsened convergence for the level-dependent scheme due the significant difference between the fine- and coarsegrid operator definition. This in particular implies that the smoothest mode weight  $|1 - e^{-id\theta}|$  becomes large, see (34), causing the additional coarsegrid correction error  $\varepsilon^{2h}$  to be distintively non-zero and even contain a significant contribution of smooth modes. This however corrupts the correction of the finegrid error causing convergence to deteriorate. Note that our standard choice for the maximal rotation of the coarsest level  $\theta_{max} = \pi/6$  seems to be rather performant for the problem given. However, the optimal choice of the perturbation parameter  $d\theta$  with respect to convergence is obviously problem-dependent.

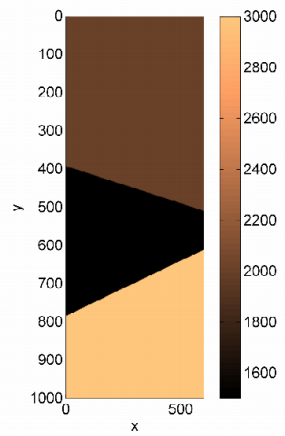
#### 4.2. The wedge model

The wedge model was introduced in [25] for the analysis of a preconditioner based on separation of variables and adopted in [19] to test the CSL preconditioner. It simulates a seismic scattering problem where a radar signal with frequency  $f \in (10Hz, 50Hz)$  is sent into the earth's surface that consists of three different layers. In each of these layers the sound waves travel at a different speed, as illustrated in Figure 4,

$$c(x, y) = \begin{cases} 2000, & (\text{if } 0 < y < \frac{1}{6}x + 400); \\ 1500, & (\text{if } \frac{1}{6}x + 400 \leq y < -\frac{1}{3}x + 800); \\ 3000, & (\text{if } -\frac{1}{3}x + 800 \leq y < 1000), \end{cases}$$

which results into a mildly space-dependent wave number given by

$$k(x, y) = \left( \frac{2\pi f}{c(x, y)} \right)^2.$$

Figure 4. Velocity profile  $c(x, y)$  for the 2D wedge problem.

$f$	10	20	30	40	50
$n_x \times n_y$	$64 \times 128$	$128 \times 256$	$128 \times 256$	$256 \times 512$	$256 \times 512$
MG-FGMRES	30 (2.22)	51 (15.3)	78 (26.6)	94 (218)	114 (294)
MG-FGMRES(10)	32 (2.10)	58 (13.8)	85 (20.3)	107 (128)	133 (159)
LVL-MG	30 (2.20)	47 (10.3)	72 (15.2)	83 (79.5)	101 (96.8)

Table VI. 2D Wedge problem. Table comparing iterations and total CPU time (in s.) for standard CSG-preconditioned FGMRES and level-dependent multigrid. A V(1,1)-cycle with GMRES(3) smoother is used for both preconditioning and level-dependent multigrid. Wavenumber-dependent discretisation following the  $kh < 0.625$  criterion for a minimum of 10 gridpoints per wavelength.

Additionally, the following parameters are used:

$$\begin{aligned}
 \Omega &= (0, 600) \times (0, 1000), \\
 \chi(x, y) &= \begin{cases} 1, & \text{for } x = 300, y = 0, \\ 0, & \text{elsewhere,} \end{cases} \\
 u &= \text{outgoing on } \partial\Omega.
 \end{aligned}$$

Table VI compares convergence results of the MG-FGMRES and LVL-MG methods on the 2D wedge problem. The observations are analogous to those made on the constant- $k$  problem: scalability with respect to the wavenumber  $k$  in terms of iterations is comparable for both methods; however the new LVL-MG solver clearly outperforms the Krylov method in terms of operation count, which is reflected in the overall CPU time. Note that restarting the outer Krylov method significantly improves CPU times over regular non-restarted FGMRES due to the limited cost of the orthogonalisation process. However, restarted FGMRES clearly has worsened  $k$ -scalability in the number of iterations. Overall, one observes that the computational cost of the LVL-MG is considerably lower than that of the preconditioned GMRES methods (cfr. CPU timings).

In Table VII the wedge problem is extended to a 3D setting. For simplicity, the third dimension (variable  $z$ ) is chosen to be an identical copy of the  $x$ -dimension from the 2D formulation. Additionally, computational time is limited by considering a moderately fine  $64 \times 128 \times 64$  grid at solution level for all wavenumbers  $k \in [10, 20]$ . Conclusions of the 2D problem are carried over to the 3D formulation, as iteration numbers are linearly rising in function of  $k$  for both solvers.

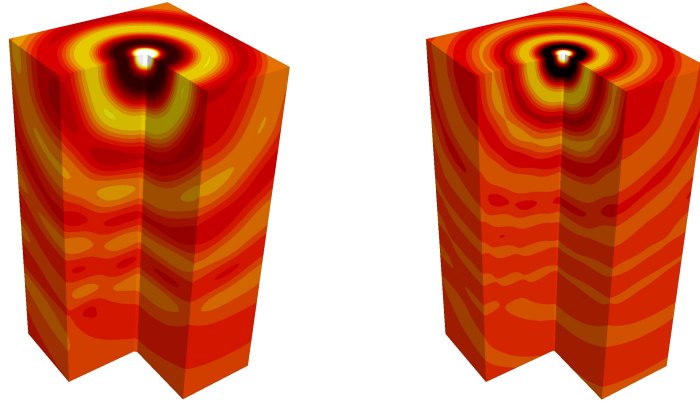


Figure 5. Real part of the solution  $u(x, y, z)$  to the 3D wedge model problem for wavenumbers  $k = 10$  (left) and  $k = 16$  (right) respectively.

$f$	12	14	16	18	20
$n_x \times n_y \times n_z$	$64 \times 128 \times 64$				
MG-FGMRES	34 (294)	39 (354)	45 (430)	53 (541)	60 (645)
MG-FGMRES(10)	35 (234)	40 (269)	47 (315)	55 (369)	62 (416)
LVL-MG	38 (191)	46 (232)	50 (252)	58 (293)	71 (356)

Table VII. 3D Wedge problem. Table comparing iterations and total CPU time (in s.) for standard CSG-preconditioned FGMRES and level-dependent multigrid. A V(1,1)-cycle with GMRES(3) smoother is used for both preconditioning and level-dependent multigrid.

Note that the CPU time gain from using LVL-MG increases even further in higher dimensions<sup>††</sup>, supporting the use of the directly applicable level-dependent multigrid method. The resulting 3D solution  $u(x, y, z)$  for wavenumbers  $k = 10$  and  $k = 16$  is plotted in Figure 5.

#### 4.3. The quantum mechanical ionization model

The Helmholtz equation can also be used to understand and predict the reaction rates of fundamental processes in few-body physics and chemistry that are of profound importance for many areas of technology. Therefore it is necessary to solve the multi-dimensional Schrödinger equation, equivalent to a multi-dimensional Helmholtz equation

$$-\Delta u(\mathbf{x}) - k(\mathbf{x})u(\mathbf{x}) = \chi(\mathbf{x}), \quad \mathbf{x} \in \Omega \subset \mathbb{R}^n, \quad (40)$$

with outgoing waves boundary conditions, source function  $\chi(\mathbf{x})$  and a space-dependent wavenumber  $k(\mathbf{x})$ .

The below 2D benchmark originates from the dynamics of two electrons in a Hydrogen molecule. Coordinates  $x$  and  $y$  should be interpreted as radial distances of particles to the center of mass of the system. Along the lines  $x = 0$  and  $y = 0$  the wave function  $u(x, y)$  is zero, while at the other sides

<sup>††</sup>The growing CPU time discrepancy between FGMRES and LVL-MG in function of the problem dimension is a logical consequence of the additional matvec execution in the outer Krylov method, which indeed gets more computationally expensive as dimension grows.

$k_0$	1	2	3	4	5
$n_x \times n_y$	$128^2$	$256^2$	$256^2$	$512^2$	$512^2$
MG-FGMRES	51 (5.40)	92 (46.8)	191 (137)	174 (813)	306 (2185)
MG-FGMRES(10)	66 (5.52)	140 (46.7)	245 (81.3)	250 (421)	393 (661)
MG-FGMRES(30)	52 (4.86)	93 (34.7)	226 (83.2)	278 (565)	426 (867)
LVL-MG	44 (3.42)	83 (25.3)	208 (63.5)	149 (215)	289 (418)

Table VIII. 2D Ionization problem. Table comparing iterations and total CPU time (in s.) for standard CSG-preconditioned FGMRES and level-dependent multigrid. A V(1,1)-cycle with GMRES(3) smoother is used for both preconditioning and level-dependent multigrid. Wavenumber-dependent discretisation following the  $kh < 0.625$  criterion for a minimum of 10 gridpoints per wavelength.

of the domain outgoing wave conditions are needed. We use the following model specifications:

$$\begin{aligned}
\Omega &= (0, r)^2, \\
k(x, y) &= \frac{1}{e^{x^2}} + \frac{1}{e^{y^2}} + k_0^2, \\
\chi(x, y) &= \frac{1}{e^{x^2+y^2}}, \\
u(0, y) = u(x, 0) &= 0, \quad (\text{Dirichlet conditions}), \\
u(r, y \neq 0), u(x \neq 0, r) &= \text{ECS}, \quad (\text{outgoing wave conditions}),
\end{aligned}$$

with  $0 < k_0 < 5$ . The domain  $\Omega$  is chosen to range from 0 to  $r = 50$  in each spatial dimension.

Serving as our final testcase, quantum mechanical ionization problems are generally considered rather hard-to-solve Helmholtz problems due to the heterogeneity in the domain. The wavenumber function  $k(x, y)$  is clearly heavily space-dependent, rendering these type of Helmholtz problems one of the major challenges in the development of efficient and robust solvers.

Convergence results on the solution of the ionization problem using both MG-FGMRES and the LVL-MG solver are displayed in Table VIII. The  $k$ -scalability is almost identical for both methods in terms of iteration count, as was observed in previous experiments. However, due to the elimination of the outer Krylov method when employing the level-dependent multigrid scheme, a speed-up of at least 30% can be achieved by direct application of the LVL-MG method. Indeed, the LVL-MG method has a drastically reduced flop count in comparison to the MG-FGMRES solver, resulting in a significantly improved CPU-time scalability in function of the wavenumber  $k_0$ .

## 5. CONCLUSIONS

In this paper, we have developed a level-dependent correction scheme for indefinite Helmholtz problems. This new level-dependent correction scheme is based on the idea of perturbing the Helmholtz operator, which was originally introduced by the Complex Shifted Laplacian and Complex Stretched Grid preconditioning techniques. These well-known schemes respectively shift or rotate the Helmholtz operator spectrum in the complex plane away from the origin, hence guaranteeing multigrid stability. The level-dependent correction scheme incorporates this idea, gradually shifting/rotating the spectrum throughout the multigrid hierarchy, thus resolving the well-known two-grid instability issues of the standard two-grid scheme while maintaining the original Helmholtz operator on the finest grid. This results in a multigrid scheme which, contrarily to CSL or CSG, is capable of *directly* solving the Helmholtz system instead of being used as a preconditioner.

The aforementioned methodology of gradually perturbing the original Helmholtz operator throughout the hierarchy introduces a new complication in ways of an additional unwanted error being added to the coarsegrid correction. This additional error is intrinsic to the level-dependent solver, as it appears as an artifact of the difference between the coarse- and finegrid residual equations. However, it is proven that this ‘perturbation error’ consist primarily of oscillatory

eigenmodes, which are subsequently damped by the action of the smoother. Consequently, given a sufficiently performant smoother, the level-dependent two-grid scheme is expected to be a very efficient and stable solver for Helmholtz problems.

To ensure effective smoothing we propose the use of GMRES(3) as a smoother substitute, which is recently shown to be highly efficient when used as a relaxation scheme in a multigrid setting. Numerical experiments on 2D and 3D Helmholtz benchmark problems of various difficulty confirm the efficiency of the proposed level-dependent correction scheme as a solver, showing it to be competitive with the current state-of-the-art CSL- or CSL-preconditioned Krylov methods. Additionally, like any multigrid solver the level-dependent scheme is shown to be fully  $h$ -independent. However, perfect  $k$ -scalability is generally not guaranteed; indeed, scalability in function of the wavenumber is largely comparable to the present-day preconditioned Krylov methods.

## 6. ACKNOWLEDGEMENTS

This research has been funded by the *Fonds voor Wetenschappelijk Onderzoek (FWO)* project G.0.120.08 and *Krediet aan navorser* project number 1.5.145.10. Additionally, this work is partly funded by Intel® and by the *Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT)*. The authors would like to thank Hisham bin Zubair for sharing a multigrid implementation.

## REFERENCES

1. J. Aguilar and J.M. Combes. A class of analytic perturbations for one-body Schrödinger Hamiltonians. *Communications in Mathematical Physics*, 22(4):269–279, 1971.
2. T. Airaksinen, E. Heikkola, A. Pennanen, and J. Toivanen. An algebraic multigrid based shifted-Laplacian preconditioner for the Helmholtz equation. *Journal of Computational Physics*, 226(1):1196–1210, 2009.
3. R.E. Bank and C.C. Douglas. Sharp estimates for multigrid rates of convergence with general smoothing and acceleration. *SIAM journal on numerical analysis*, pages 617–633, 1985.
4. A. Bayliss, C.I. Goldstein, and E. Turkel. An iterative method for the helmholtz equation. *Journal of Computational Physics*, 49(3):443–457, 1983.
5. A. Bayliss, C.I. Goldstein, and E. Turkel. On accuracy conditions for the numerical computation of waves. *Journal of Computational Physics*, 59(3):396–404, 1985.
6. J.P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *Journal of computational physics*, 114(2):185–200, 1994.
7. M. Bollhöfer, M.J. Grote, and O. Schenk. Algebraic multilevel preconditioner for the Helmholtz equation in heterogeneous media. *SIAM Journal on Scientific Computing*, 31(5):3781–3805, 2009.
8. A. Brandt. Multi-level adaptive solutions to boundary-value problems. *Mathematics of computation*, 31(138):333–390, 1977.
9. A. Brandt and I. Livshits. Wave-ray multigrid method for standing wave equations. *Electronic Transactions on Numerical Analysis*, 6:162–181, 1997.
10. A. Brandt and S. Ta’asan. Multigrid method for nearly singular and slightly indefinite problems. *Multigrid Methods II, Lecture Notes in Math.*, 1228:99–121, 1986.
11. W.L. Briggs, V.E. Henson, and S.F. McCormick. *A multigrid tutorial*. Society for Industrial Mathematics, Philadelphia, 2000.
12. H. Calandra, S. Gratton, R. Lago, X. Pinel, and X. Vasseur. Two-level preconditioned Krylov subspace methods for the solution of three-dimensional heterogeneous Helmholtz problems in seismics. Technical report, TR/PA/11/80, CERFACS, Toulouse, France, 2011.3, 2011.
13. S. Cools and W. Vanroose. Local Fourier Analysis of the Complex Shifted Laplacian preconditioner for Helmholtz problems. *Numerical Linear Algebra with Applications*, (submitted), 2012.
14. H.C. Elman, O.G. Ernst, and D.P. O’Leary. A multigrid method enhanced by Krylov subspace iteration for discrete Helmholtz equations. *SIAM Journal on scientific computing*, 23(4):1291–1315, 2002.
15. Y.A. Erlangga and R. Nabben. On a multilevel krylov method for the helmholtz equation preconditioned by shifted Laplacian. *Electronic Transactions on Numerical Analysis*, 31(403–424):3, 2008.
16. Y.A. Erlangga and R. Nabben. Algebraic multilevel krylov methods. *SIAM Journal on Scientific Computing*, 31:3417–3437, 2009.
17. Y.A. Erlangga, C.W. Oosterlee, and C. Vuik. On a class of preconditioners for solving the Helmholtz equation. *Applied Numerical Mathematics*, 50(3–4):409–425, 2004.
18. Y.A. Erlangga, C.W. Oosterlee, and C. Vuik. A novel multigrid based preconditioner for heterogeneous Helmholtz problems. *SIAM Journal on Scientific Computing*, 27(4):1471–1492, 2006.

19. Y.A. Erlangga, C. Vuik, and C.W. Oosterlee. Comparison of multigrid and incomplete LU shifted-Laplace preconditioners for the inhomogeneous Helmholtz equation. *Applied Numerical Mathematics*, 56(5):648–666, 2006.
20. O.G. Ernst and M.J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In *Numerical Analysis of Multiscale Problems. Durham LMS Symposium*. Citeseer, 2010.
21. E. Haber and S. MacLachlan. A fast method for the solution of the Helmholtz equation. *Journal of Computational Physics*, 230(12):4403–4418, 2011.
22. A. Laird and M. Giles. Preconditioned iterative solution of the 2d Helmholtz equation. Technical report, NA-02/12, Comp. Lab. Oxford University UK, 2002.
23. S.P. MacLachlan and C.W. Oosterlee. Algebraic multigrid solvers for complex-valued matrices. *SIAM Journal on scientific computing*, 30(3):1548–1571, 2008.
24. D. Osei-Kuffuor and Y. Saad. Preconditioning Helmholtz linear systems. *Applied Numerical Mathematics*, 60(4):420–431, 2010.
25. R.E. Plessix and W.A. Mulder. Separation-of-variables as a preconditioner for an iterative Helmholtz solver. *Applied Numerical Mathematics*, 44(3):385–400, 2003.
26. B. Reps and W. Vanroose. Analyzing the wave number dependency of the convergence rate of a multigrid preconditioned Krylov method for the Helmholtz equation with an absorbing layer. *Numerical Linear Algebra with Applications*, 19(2):232–252, 2012.
27. B. Reps, W. Vanroose, and H. bin Zubair. On the indefinite Helmholtz equation: Complex stretched absorbing boundary layers, iterative analysis, and preconditioning. *Journal of Computational Physics*, 229(22):8384–8405, 2010.
28. B. Reps, W. Vanroose, and H. bin Zubair. Multigrid preconditioners based on polynomial smoothers for the Helmholtz equation with absorbing layers. *Submitted*, 2012.
29. C.D. Riyanti, Y.A. Erlangga, R.E. Plessix, W.A. Mulder, C. Vuik, and C. Oosterlee. A new iterative solver for the time-harmonic wave equation. *Geophysics*, 71(5):E57–E63, 2006.
30. Y. Saad. FGMRES: A flexible inner-outer preconditioned GMRES algorithm. *SIAM Journal on Scientific Computing*, 14(3):461–469, 1993.
31. Y. Saad and M.H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7:856–869, 1986.
32. A.H. Sheikh, D. Lahaye, and C. Vuik. A scalable Helmholtz solver combining the shifted Laplace preconditioner with multigrid deflation. Technical report, DIAM 11-01, Delft University of Technology, 2011.
33. V. Simoncini and D.B. Szyld. Flexible inner-outer Krylov subspace methods. *SIAM Journal on Numerical Analysis*, pages 2219–2239, 2003.
34. K. Stüben and U. Trottenberg. Multigrid methods: Fundamental algorithms, model problem analysis and applications. *Multigrid methods, Lecture Notes in Math.*, 960:1–176, 1982.
35. U. Trottenberg, C.W. Oosterlee, and A. Schüller. *Multigrid*. Academic Press, New York, 2001.
36. H.A. van Gijzen, Y. Erlangga, and C. Vuik. Spectral analysis of the discrete Helmholtz operator preconditioned with a shifted Laplacian. *SIAM Journal on Scientific Computing*, 29(5):1942–1985, 2007.
37. R.S. Varga. *Matrix iterative analysis*. Englewood Cliffs, N.J., Prentice-Hall, 2 edition, 1962.